Centre de recherche sur les transports (C.R.T.)
Centre for Research on Transportation

CRT-2005-10

CAPACITY CONSIDERATIONS IN TRANSIT ASSIGNMENT: A REVIEW

by

Younes Hamdouch
Michael Florian

April 2005

Université de Montréal
C.P. 6128, succursale Centre-ville, Montréal QC H3C 3J7 Canada
Téléphone : (514) 343-7575 / Télécopieur : (514) 343-7121

crt@crt.umontreal.ca
www.crt.umontreal.ca

HEC ÉCOLE POLYTECHNIQUE MONTRÉAL Université de Montréal

105-05371/

PUBLICATION
CRT-2005-10

CAPACITY CONSIDERATIONS IN TRANSIT ASSIGNMENT: A REVIEW

by

Younes Hamdouch [1]
Michael Florian [2]

The research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

[1] Business Administration Department, United Arab Emirates University, Al Ain, U.A.E.

[2] Département d'informatique et de recherche opérationnelle and Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Canada H3C 3J7 (mike@crt.umontreal.ca)

Dépôt légal – Bibliothèque nationale du Québec
Bibliothèque nationale du Canada, 2005

ABSTRACT

This paper reviews the main contributions to the consideration of capacities in transit route choice models. They are based on transit route choice models without capacity, which are all reviewed in detail.

Keywords: transit networks, equilibrium assignment, strategy.

RÉSUMÉ

Cet article présente les principales contributions à la considération des capacités dans les modèles d'affectation de transport en commun. Ces approches sont basées sur des modèles sans capacité qui sont décrits en détail.

Mots-Clés: transport en commun, affectation d'équilibre, stratégie.

1 Introduction

Transit assignment models aim to represent the passenger behaviour in public transportation networks in order to identify the paths used, given that origin-destination travel demands and transit line parameters (capacities, frequencies and headway distributions) are known. Several models have been proposed in the literature, differing with respect to the assumptions on passenger behaviour, network representation and modeling of congestion.

In the seventies, the transit assignment problem has been studies, either as a separate problem (Dial, 1967; Fearnside and Draper, 1971; Chriqui and Robillard, 1975; Chapleau, 1974; Rapp *et al.*, 1976; Andreasson, 1976) or as a subproblem of transit network design (Lampkin and Saalmans, 1967; Mandl, 1980; Hasselström, 1981). With the exception of the model of Fearnside and Draper (1971) where passenger flows are assigned to the *shortest single path*, most modelling approaches that have been employed are based on the fundamental concept of *strategy*. This idea states that, since the intervals separating passenger arrivals from bus arrivals at transit stops are random, then it could be advantageous for passengers to select a set of attractive lines at each reached bus stop (rather than a single line) and board the first arriving bus. For example, in the model of Dial (1967), the idea of strategy is restricted to the case where only *common* lines with equivalent speeds can be selected in the set of *attractive lines* at bus stops. Dial proposed to replace the sets of common lines by super-lines (or *truck-lines*) bearing a frequency equal to the sum of frequencies of each individual lines. Le Clercq (1972) adopted a similar approach but based on a line-node description of the transit network. With the difference of Dial, Le Clercq supposed that all common lines are different and therefore passengers are assigned to each one of them proportionally to their frequency. Chriqui and Robillard (1975) presented a more general common line model where transit lines can have different travel times. If there are several common lines between two consecutive stops, passengers will select a subset of attractive lines that minimizes the total expected time. In the TRANSCOM model (Chapleau, 1974) which used the algorithm known as CTM, the existence of common lines is not considered and all passengers travelling between a given pair of nodes used *transit routes* to reach their destination. Adopting the same approach of transit route, De Cea and Fernández (1989) proposed a mathematical formulation of the transit assignment problem where passengers travel by following a sequence of intermediate transfer nodes. De Cea and Fernández (1993) proposed a similar approach to congested networks where waiting times as well as passenger

distribution depend not only on the flow boarding at the node but also on the flow already on the bus.

As an example, consider the small transit network of Figure 1. An example of a *single path* is illustrated in Figure 2. This single path can be denoted: *AD*(*l*1)-*DF*(*l*2). There are 7 single paths to travel form *A* to *F*. The path already mentioned and the 6 following: *AB*(*l*4)-*BF*(*l*4), *AC*(*l*3)-*CF*(*l*3), *AD*(*l*1)-*DF*(*l*1), *AD*(*l*2)-*DF*(*l*1), *AD*(*l*2)-*DF*(*l*2) and *AE*(*l*5)-*EF*(*l*6). The model of Fearnside and Draper permits to the users to choose among these 7 itineraries. An example of a path with common lines is illustrated in Figure 3. We denoted this superpath by *AD*(*l*1, *l*2)-*DF*(*l*1,*l*2), If we adopt the common line model of Chriqui and Robillard, we obtain 12 paths with common lines to reach node *F* from node *A*. In addition to the 7 simple paths (model of Fearnside and Draper) there exists 5 itineraries using the set of attractive lines (*l*1, *l*2) between *A* and *F*: *AD*(*l*1, *l*2)-*DF*(*l*1), *AD*(*l*1, *l*2)-*DF*(*l*2), *AD*(*l*1)-*DF*(*l*1,*l*2), *AD*(*l*2)-*DF*(*l*1,*l*2), and *AD*(*l*1, *l*2)-*DF*(*l*1, *l*2). Finally, Figure 4 show an example of a transit route *A* (*l*1, *l*2, *l*3, *l*4)-*B*(*l*4)-*C*(*l*3)-*D*(*l*1,*l*2).

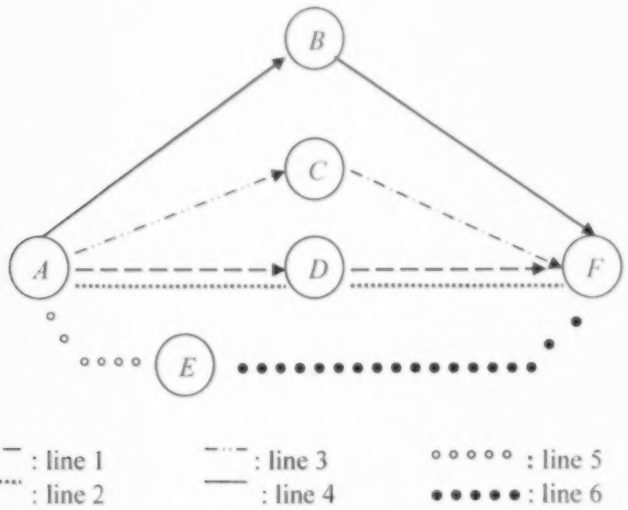

Figure 1: Small network

Spiess (1983), (see also Spiess and Florian, 1989) were the first to introduce the idea of *strategy* for the transit traveler's route choice behaviour. A strategy is a set of rules that, when applied, allows users to choose a route from their origin to their destination. The king of different strategies that travelers may choose from depends on the information that is available to them during the trip. In the model of Spiess, the only information available to users is, while waiting at node, which is the next line to be served, i.e. users may at that time

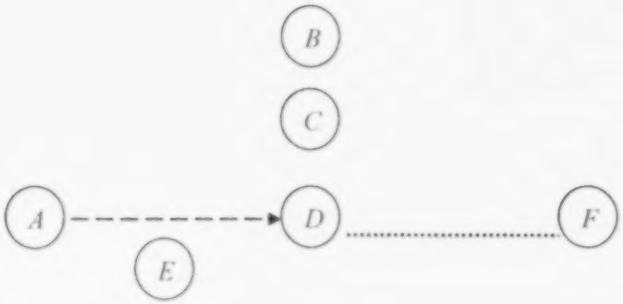

Figure 2: Example of a single path

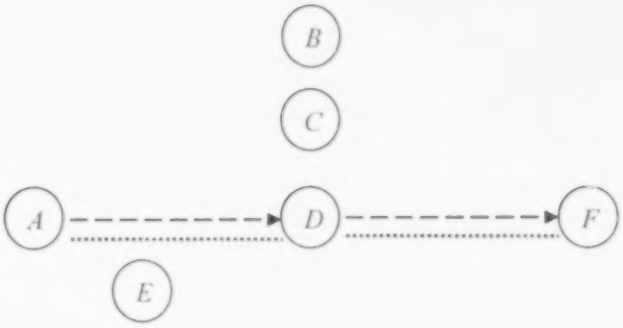

Figure 3: Example of a path with common lines

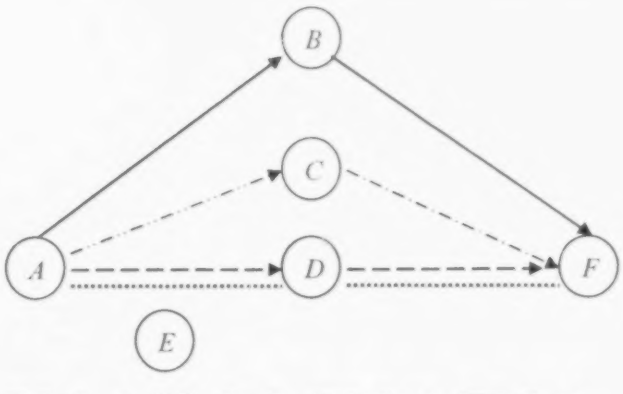

Figure 4: An example of a transit route

decide whether or not to board this transit bus with the aim to minimise the expected waiting and travel time. It is then possible to define a strategy of this type by specifying at every node of the network, a set of attractive lines. In the example of Figure 1, a possible strategy for passengers travelling from the origin node *A* to the destination node *F* is: *A*(*l*1,*l*2,*l*3,*l*5)-*C*(*l*3)-*D*(*l*1,*l*2)-*E*(*l*6) (see Figure 5). Spiess and Florian (1989) proposed a linear formulation of the strategic transit assignment problem and proposed a label-setting algorithm to solve the problem. Gendreau (1984) developed a similar model for congested networks in which the waiting process at bus stops was based on a bulk queue approach. Spiess and Florian (1989) proposed a non-linear extension for their linear model where link travel times are continuous functions of the vector of link flows. Another main contribution has been brought by Nguyen and Pallottino (1988) who characterized a strategy as an acyclic directed graph for each destination, and denoted it as a *hyperpath*. The authors developed an equilibrium assignment model where travel times are flow-dependent and passengers travel along shortest hyperpaths. Wu and Florian (1993), Wu, Florian and Marcotte (1994) extended the last models by considering both travel times and waiting times to be flow dependent. The authors proposed several solution algorithms, based on the linearized Jacobi method, the projection method and the simplicial decomposition strategy. Bouzaïene-Ayari (1996) and Bouzaïene-Ayari, Gendreau and Nguyen (1995) proposed a congested assignment model where all travel costs/times and distribution probabilities at bus stops are flow-dependent. The distribution of transit flows among attractive lines is done in inverse proportion of the waiting times, which may be flow dependent.

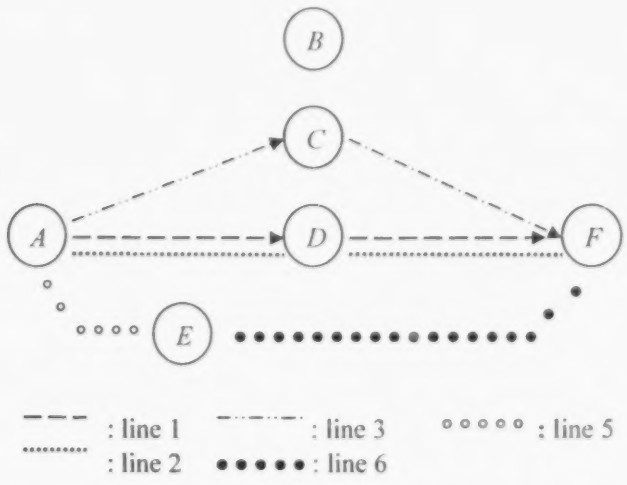

Figure 5: An example of a strategy

The above mentioned studies considered only congestion aboard the vehicles. However, if the arriving vehicle has limited capacity, then not all passengers waiting at a stop can board the first arriving vehicle and part of these passengers have to wait for one of the following vehicles would have available seats in order to board. Cominetti and Correa (2001) were the first to take into account this effect. The authors developed a congested common-line model in which the effective frequency of transit lines is a strictly decreasing function of total flows and congestion affects both the waiting times and the flow distribution. Cepeda (2002) extended this model to general transit networks by using a *node-strategy* approach that permits a new formulation of the assignment problem. He proposed a new characterization of the route choice equilibria based exclusively on the arc-flows. Marcotte and Nguyen (1998) and Hamdouch (2002) studied the transit assignment problem in networks involving arcs with rigid capacities where the sole congestion effects are induced by the link capacities. Contrary to the notion of strategy in standard transit models (Spiess and Florian, 1989; Cominetti and Correa, 2001) where randomness is driven by frequencies, the authors used a different definition of strategy where the ordering of attractive lines matters. In this case, a strategy associates with each node of the network an ordered set of successor nodes and a passenger selects the first available outgoing arc in this ordered list, where the probability that an arc be available is proportional to its capacity and inversely proportional to the number of users longing for it. The authors developed an algorithmic framework and provided numerical results on test networks.

In this paper, we review several mathematical approaches that have been proposed for the passenger assignment problem. First, we study in Section 2 transit assignment models that consider the travel times and the frequencies to be constant. Then, Section 3 is devoted to transit models where congestion affects the travel time and waiting time at bus stops. Next, we survey in Section 4 fully-congested models that considers both congestion aboard the vehicles and the limited capacity of transit vehicles. Finally, we conclude the paper with a discussion of several open questions and future research directions.

2 Transit assignment model-linear travel costs and uncapacitated

Consider a transit network consisting of a set of transit lines and walk link where passengers board and alight vehicles. At a given stop, a passenger may have the choice of several lines

and itineraries to reach his destination. In this part of the review, the time associated with each walk line and each transit line segment is taken to be constant.

2.1 Model of Chriqui and Robillard – Common lines

Chriqui and Robillard (1975) formulated and solved the common lines problem on a simple network of one origin A and one destination B where passengers can select a subset of attractive lines that minimizes the expected waiting and travel times. Let L be the set of all transit lines travelling from node A and going to node B. Chriqui and Robillard assume that a passenger willing to travel from A to B will only consider a subset of lines $\overline{L} \subset L$ in order to minimise his expected total travel time. Given this selection, he will board the first vehicle which arrives to node A and belongs to any line $l \in \overline{L}$. The total expected time between node A and B is then:

$$TV_{\overline{L}} = W_{\overline{L}} + T_{\overline{L}}, \tag{1}$$

where $W_{\overline{L}}$ and $T_{\overline{L}}$ represent the expected waiting time and in-vehicle travel time respectively. All lines l belonging to $\overline{L}$ are named *common* or *attractive* lines to travel from A and B. If f_l and t_l denote the frequency and in-vehicle travel time associated to the line $l \in L$ and the waiting times for these lines are supposed to be independent random variables with means $1/f_1, 1/f_2, \cdots, 1/f_n$, the solution of the following problem defines the set of common lines L_2:

$$\min_{\{x_l\}} \quad \frac{1 + \sum_{l \in L} t_l f_l x_l}{\sum_{l \in L} f_l x_l} \tag{2}$$

$$\text{subject to} \quad x_l = 0,1 \quad \forall l \in L,$$

where the objective function denotes the total expected travel time given that :

$$W_{\overline{L}} = \frac{1}{\sum_{l \in L} f_l x_l}$$

$$T_{\overline{L}} = \frac{\sum_{l \in L} t_l f_l x_l}{\sum_{l \in L} f_l x_l}$$

A line $l \in L$ is part of the common lines set if $x_l = 1$ and it will not be used to travel between A and B if $x_l = 0$. Chriqui and Robillard developed a linear-time algorithm for computing the

set of attractive lines $\overline{L}$: initialize $\overline{L} \leftarrow \left\{ l : t_l = \min_k t_k \right\}$ and repeat the update $\overline{L} \leftarrow \overline{L} \cup \left\{ l \notin \overline{L} : t_l < TV_{\overline{L}} \right\}$ until no line is added.

2.2 Model of Spiess and Florian – Strategies

Spiess and Florian (1989) developed a model for the transit assignment problem where passengers choose *strategies* that allow them to reach their destination at minimum expected cost. Let $G = (N, A)$ denote a transit network consisting of a set of nodes N and a set of directed arcs A. The demand for travel from nodes $i, i \in N$ to the destination q is denoted g_i. Let $A_i^+ (A_i^-)$ be the set of outgoing (incoming) arcs from (to) node i. A strategy $\overline{A}$ is a mapping that's associates with every node $i, i \in N$ a set of *attractive lines* A_i^+ at node i. At each node $i, i \in N$, a passenger adopting strategy $\overline{A}$ boards the first vehicle that serves any of the lines in the set $\overline{A}_i^+ \subset A_i^+$. Let $W(\overline{A}_i^+)$ be the expected waiting time for the arrival of the first vehicle serving any of the links $a \in \overline{A}_i^+$ and $P_a\left(\overline{A}_i^+\right)$ be the probability that link a is the first to be served among the links of $\overline{A}_i^+$. If an exponential distribution of interarrival times is assumed then

$$W\left(\overline{A}_i^+\right) = \frac{1}{\sum_{a \in \overline{A}_i^+} f_a} \tag{3}$$

$$P\left(\overline{A}_i^+\right) = \frac{f_a}{\sum_{a \in \overline{A}_i^+} f_a}, \quad a \in \overline{A}_i^+, \tag{4}$$

where f_a is the frequency of link (line) a.

Since the strategy $\overline{A}$ is not known a priori, we need to express $\overline{A}$ in terms of the 0-1 variables x_a

$$x_a = \begin{cases} 0 & \text{if} \quad a \notin \overline{A} \\ 1 & \text{if} \quad a \in \overline{A} \end{cases}, \quad a \in A. \tag{5}$$

The problem of finding the optimal strategy $\overline{A}^*$ may be stated as follows:

$$\min \sum_{a \in A} s_a v_a + \sum_{i \in N} \frac{V_i}{\sum_{a \in \overline{A}_i^+} f_a x_a} \tag{6}$$

$$\text{subject to } v_a = \frac{x_a f_a}{\sum_{a' \in A_i^+} f_{a'} x_{a'}}, \quad a \in A_i^+, i \in N \tag{7}$$

$$V_i = \sum_{a \in A_i^-} v_a + g_i, \quad i \in N \tag{8}$$

$$V_i \geq 0, \quad i \in N \tag{9}$$

$$x_a = 0 \text{ or } 1, \quad a \in A, \tag{10}$$

where s_a is the travel cost on link a and V_i is the total volume at node i. At first sight, the problem (7)-(10) is a mixed integer nonlinear optimization problem. However, Spiess and Florian proved that this problem may be reduced to the following simpler linear programming problem:

$$\min \sum_{a \in A} s_a v_a + \sum_{i \in N} w_i \tag{11}$$

$$\text{subject to } \sum_{a \in A_i^+} v_a - \sum_{a \in A_i^-} v_a = g_i, \quad i \in N \tag{12}$$

$$v_a \leq f_a w_i, \quad a \in A_i^+, \ i \in N \tag{13}$$

$$v_a \geq 0, \quad a \in A, \tag{14}$$

where $w_i = \dfrac{V_i}{\sum_{a \in A_i^+} f_a x_a}$, $i \in N$ denotes the total waiting time of all trips at node i.

Spiess developed a solution algorithm to solve the transit route choice model corresponding to the primal and dual formulations of the problem (11)-(15). This algorithm bears a strong resemblance to the label-setting algorithm for computing shortest paths:

Part 1: Find optimal strategy

1.1	(Initialisation)	$u_i = \infty, i \in N - \{r\} := 0$; expected total travel times $f_i := 0, i \in N$; $S := A; \overline{A} := \theta$.
1.2	(Get next link)	If $S = \theta$ then STOP Otherwise find $a = (i, j) \in S$ which satisfies $u_j + c_a \leq u_{j'} + c_{a'}, a' = (i', j') \in S$; $S := S - \{a\}$.
1.3	(Update node label)	If $u_i \geq u_j + c_a$ then

$$u_i := \frac{f_i u_i + f_a (u_i + c_a)}{f_i + f_a}$$

$f_i := f_i + f_a$, $\overline{A} := \overline{A} + \{a\}$;

go to step 1.2.

Part 2: Assign demand according to optimal strategy

2.1 (Initialisation) $V_i := g_i, i \in N$;

2.2 (Loading) Do for every link $a \in A$, in decreasing order of $(u_j + c_a)$;

if $a \in \overline{A}$ then $v_a := \frac{f_a}{f_i} V_i$

$V_j := V_j + v_a$,

otherwise $v_a := 0$.

2.3 Model of Nguyen and Pallottino – Hyperpaths

Nguyen and Pallottino (1988) provided a graph theoretic interpretation of a strategy (introduced by Spiess, 1983) as an acyclic directed graph for each destination, and denoted it as a *hyperpath*. Consider a transit network $G = (N, E)$ consisting of a set of nodes N and a set of arcs E. An hyperpath k connecting an origin r to destination s is as subgraph $G_k = (N_k, E_k, \pi_k)$, where $N_k \subset N$, $E_k \subset E$ and $\pi_k = (\pi_{ijk})$ a vector of access probabilities of dimension $|E_k|$, such that :

- G_k is acyclic with at least one arc;
- node r has no predecessors and s no successors;
- for every node $i \in N_k - \{r, s\}$, there is a path from r to s traversing i, and if node $i \notin R$ (set of stops nodes) then i has at most one immediate successor;
- the vector π_k satisfies

$$\sum_j \pi_{ijk} = 1, \quad \forall i \in N_k,$$

$$\pi_{ijk} \geq 0, \quad \forall (i, j) \in E_k.$$

For every stop node i of hyperpath k, the set of boarding arcs E_{ik}^+ identifies the subset of transit lines associated with hyperpath k. Each component π_{ijk} is equal to the probability that

arc (i, j) is chosen among the arcs of E_{ik}^{+}. These arcs probabilities induce the *path probabilities*:

$$\lambda_p = \prod_{(i,j)\in E_k} \pi_{ijk}^{\delta_{ijp}}, \quad \forall p \in P_k,$$

where P_k denotes the set of all paths of k connecting r and s, δ_{ijp} equals 1 if path p traverses arc (i, j), and 0 otherwise and $\sum_{p\in P_k} \lambda_p = 1$. Note that an elementary path is a hyperpath where every node has an access probability equal to zero or one. In addition to the ordinary arc cost c_{ij}, a fixed unit waiting time w_{ik} is also associated with each stop node $i \in R_k$. The cost of any path $p \in P_k$ is simply the sum of the costs of the arcs and nodes forming path p :

$$c_p = \sum_{(i,j)\in E_k} c_{ij} + \sum_{i\in R_k} w_{ik}, \quad \forall p \in P_k.$$

The cost of a hyperpath k is defined as the weighted cost of all elementary paths contained in the hyperpath :

$$C_k = \sum_{p\in P_k} \lambda_p c_p = \sum_{p\in P_k} \lambda_p \Big(\sum_{(i,j)\in E_k} c_{ij} + \sum_{i\in R_k} w_{ik} \Big),$$

or equivalently :

$$C_k = \sum_{(i,j)\in E} \alpha_{ijk} c_{ij} + \sum_{i\in N} \beta_{ik} w_{ik},$$

where the node probabilities β_{ik} and arc probabilities α_{ijk} are defined as:

$$\beta_{jk} = \sum_{i\in N_j} \beta_{ik} \pi_{ijk}, \quad \forall j \in N_k - \{r\},$$

$$\alpha_{ijk} = \beta_{ik} \pi_{ijk}, \quad \forall (i, j) \in E_k.$$

Nguyen and Pallottino assumed that passengers arrive randomly at every stop node, and that all transit lines are independently scheduled with exponential distributions of their respective headways, with mean equal to the inverse of the corresponding line frequency. In this case, the access probabilities π_{ijk} and the mean waiting time w_{ik}, at stop node i, take on specific values:

$$\pi_{ijk} = f_{ij} / \sum_{(i,j)\in E_{ik}^{+}} f_{ij}, \quad \forall (i, j) \in B_k, \tag{15}$$

$$w_{ik} = a / \sum_{(i,j)\in E_{ik}^{+}} f_{ij}, \tag{16}$$

where f_{ij} is the mean frequency of the transit line associated with boarding arc (i, j), and a is a suitable positive parameter.

Nguyen and Pallottino developed a label-correcting procedure to compute shortest hyperpaths. This procedure is based on the *Generalized Bellman's equation*: The residual cost C_{ik^*} of a shortest hyperpath k^* from any node i to the destination s satisfies

$$C_{ik^*} = \begin{cases} 0 & \text{if } i = s \\ \min\limits_{E_{ik}^+ \subset N_i^+} \left\{ w_{ik} + \sum\limits_{j \in E_{ik}^+} \pi_{ijk} \left(c_{ij} + C_{jk^*} \right) \right\} & \forall i \in N_k - \{s\}. \end{cases}$$

Using the equations (15)-(16) the above cost expression may be rewritten as

$$C_{ik^*} = \begin{cases} 0 & \text{if } i = s \\ \min\limits_{j \in N_i^+} \{ c_{ij} + C_{jk^*} \} & \text{if } i \text{ is a non-stop node} \\ \min\limits_{E_{ik}^+ \subset N_i^+} \left\{ \left(\sum\limits_{j \in E_{ik}^+} f_{ij} C_{jk^*} + 1 \right) / \sum\limits_{j \in E_{ik}^+} f_{ij} \right\} & \text{if } i \text{ is a stop node} \end{cases}$$

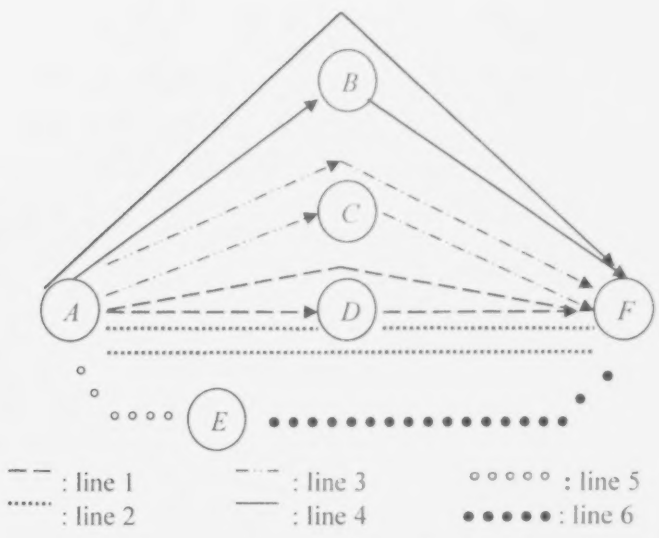

Figure 6: Graph based on line sections

The procedure can be described as follows:

(1) (Initialisation)

Set $C_{sk^*} := 0; C_{ik^*} := +\infty, \forall i \neq s$

(2) (Updating node labels)

If $C_{ik^*} > C_{jk^*} + c_{ij}$, and i is a non-stop node

then $C_{ik^*} := C_{jk^*} + c_{ij}$;

if $C_{ik^*} > C_i^* = \min_{E_{ik}^+ \subset N_i^*} \left\{ \left(\sum_{j \in E_{ik}^+} f_{ij} C_{jk^*} + 1 \right) / \sum_{j \in E_{ik}^+} f_{ij} \right\}$ and i is a stop node

then $C_{ik^*} := C_i^*$;

(3) Repeat step (2) until no label can be further decreased.

It should be noted that the model solved is the same as that of Spiess and Florian (1989).

2.4 Model of De Cea and Fernández – Transit routes

De Cea and Fernández (1989) proposed a mathematical model of the transit assignment problem to minimal routes when passengers adopt *transit routes* to travel from their origin to their destination. A transit route is any path that a user can follow on the transit network in order to travel between two nodes. The model of De Cea and Fernández is based on a rather different definition of the transit network. The authors defined a graph $G' = (N, L)$ where L is the set of all possible line sections over the network $G = (N, A)$, a line section being a sequence of transit line segments that a traveller may use without effectuating a transfer. Figure 6 shows the transit network coded in terms of nodes and lines sections over the network on figure 1. Let denote $\overline{W}$ the set of origin-destination pairs connected by at least one line section and S_{ij} the set of line sections connecting directly nodes i and j. Let $L_i^+ \left(L_i^- \right)$ be the set of outgoing (incoming) line sections from (to) node i. The total flow on route section (i, j) is denoted V_{ij}. Given these notations and the assumptions considered in Chriqui and Robillard (1975)

Arc (i,j)	Cost $C_{ij(l)}$	Arc (i,j)	Cost $C_{ij(l)}$
AD (l1)	10	CF (l3)	17
DF (l1	14	AB (l4)	17
AD (l2)	10	BF (l4)	20
DF (l2)	16	AE (l5)	7
AC (l3)	9	EF (l6)	12

Table 1: Small network: arc data

concerning the traveller's route choice process, De Cea and Fernández formulated the transit assignment to minimal routes as the following linear problem:

$$\min \sum_{l \in L} t_l v_l + \sum_{(i,j) \in \bar{W}} w_{ij} \tag{17}$$

$$\text{subject to} \quad \sum_{l \in L_i^+} v_l - \sum_{l \in L_i^-} v_l = g_i, \quad \forall i \in N \tag{18}$$

$$v_l \le f_l w_{ij}, \quad \forall (i,j) \in \bar{W}, l \in S_{ij} \tag{19}$$

$$v_l \ge 0, \quad l \in L, \tag{20}$$

where $w_{ij} = \dfrac{v_{ij}}{\sum_{l \in S_{ij}} f_l x_l}$, $\forall (i,j) \in N$, denotes the total waiting time for users travelling from node i to node j.

This linear formulation is identical to the transit assignment to optimal strategies proposed by Spiess and Florian (1989) but formulated on a different network. However, while the solution algorithm of Spiess and Florian is derived from the linear formulation of the problem, De Cea and Fernández proposed an algorithm based on the non-linear equivalent formulation when we replace the constraints (19) by the non-linear constraints ($v_l = x_l f_l w_{ij}$). De Cea *et al.* (1988) presented the transit flows obtained with these algorithms when applied to the public transit system of the city of Santiago and compared the differences of the simulated flows.

To establish the different minimal travel time between the transit route model of De Cea and Fernández and the strategic model of Spiess, let consider the small network of figure 1 where link travel times are shown in Table 2 and each line have an equal frequency of five vehicles per hour. Waiting time is assumed to be half the headway of the line, i.e. 6 minutes.

The minimal simple path for this example is $p^* = AD(l1) - DF(l1)$ whit expected cost equal to:

$$\begin{aligned} C_{p^*} &= C_{AD(l1)} + C_{DF(l1)} + w_{l1} \\ &= 10 + 14 + 6 = 30, \end{aligned}$$

where w_{lk} is the waiting time associated to line $lk (k = 1, 2, ..., 6)$.

Now, the minimal transit route is $r^* = A(C,D) - C(F) - D(F)$. $33\frac{1}{3}\%$ of passengers adopting this route use line $l1$, $33\frac{1}{3}\%$ use line $l2$ and $33\frac{1}{3}\%$ use line $l3$. The expected cost of this minimal route is

$$C_{r^*} = \frac{1}{3}(C_{AD(l1)} + C_{DF(l1)} + C_{AD(l2)} + C_{DF(l2)} + C_{AC(l3)} + C_{CF(l3)} + w_{l1} + w_{l2} + w_{l3})$$

$$= \frac{1}{3}(10 + 14 + 10 + 16 + 9 + 17 + 2 + 2 + 2) = 27\frac{1}{3}.$$

Finally, the minimum strategy is $s^* = A(C,D,E) - C(F) - D(F) - E(F)$. 25% of passengers adopting this strategy are assigned to line $l1$, 25% to line 2, 25% to line 3 and 25% to lines $l5$ and $l6$ (transferring at node E). The expected cost of this minimum strategy is

$$C_{r^*} = \frac{1}{4}(C_{AD(l1)} + C_{DF(l1)} + C_{AD(l2)} + C_{DF(l2)} + C_{AC(l3)} + C_{CF(l3)}$$

$$+ C_{AE(l5)} + C_{EF(l6)} + w_{l1} + w_{l2} + w_{l3} + w_{l5} + w_{l6})$$

$$= \frac{1}{4}(10 + 14 + 10 + 16 + 9 + 17 + 7 + 12 + 12) = 26\frac{1}{4}.$$

The results of this example show the different expected total travel times and assigned flows when using simple paths, transit routes and strategies.

3 Transit assignment model-soft capacity constraints

We present in this section transit assignment models when congestion has a direct impact on the assignment of passengers over the network. This part of the review will include only transit models where the capacities of transit vehicles have not been considered explicitly. In this case, congestion can increase the travel time and waiting time as bus stops or also modify users distribution to various transit lines.

3.1 Model of Spiess and Florian – Strategies

Spiess and Florian (1989) proposed a non-linear extension for their linear model (section 2.2) where the link travel times c_a are no longer constants, but are continuous functions $c_a(v), a \in A$, of the vector of link flows v. Such functions may model queuing delays at stops or the discomfort "cost" of passengers in crowded vehicles.

Let v_a^r denote the volume on link $a, a \in A$, that is associated with destination node $r, r \in R$. The total volume v_a on link a is the sum of the link volumes v_a^r to all destinations $r \in R$:

$$v_a = \sum_{r \in R} v_a^r, \quad a \in A. \tag{21}$$

The problem is no longer separable by destination, as is the case for the linear cost model. The demand from node $i, i \in N$ to destination $r, r \in R$, is denoted g_i^r and K_r denote the set of all feasible strategies to reach destination r. h_i^k is the part of the demand g_i^r that is assigned to the network adopting strategy $k \in K_r$. The conservation of flow equations is:

$$\sum_{k \in K_r} h_i^k = g_i^r, \quad i \in N, k \in K_r, r \in R. \tag{22}$$

In the same way that the linear model, it is assumed that travellers choose strategies that minimize the expected travel times of their trips. However, in the non-linear model, the travel times depend on the vector of link flows v and therefore on the strategies chosen by the other travellers. In this case, a user optimal traveler's behaviour can be characterized by equilibrium conditions derived from Wardrop's (1952) second user optimal principle. If s_i^k denotes the expected travel time of strategy $k \in K_r$ from node i to destination r, and u_i^r denotes the minimum expected travel time from i to r, the equilibrium conditions may be stated as follows:

$$\left.\begin{array}{ll} s_i^{k^*} = u_i^{r^*} & \text{if} \quad h_i^{k^*} > 0 \\ s_i^{k^*} \geq u_i^{r^*} & \text{if} \quad h_i^{k^*} = 0 \end{array}\right\}, \quad i \in N, k \in K_r, r \in R. \tag{23}$$

By using a similar development to that used by Smith (1979), it may be restated in the form of a variational inequality. For general costs $c_a(v)$ this variational inequality may not have a unique solution since the corresponding cost mapping is not necessarily monotone. When the link travel times $c_a(v)$ are continuous increasing functions $c_a(v_a)$ of the total volume on link a, there exists the following convex cost differentiable equivalent optimization problem:

$$\min \sum_{a \in A} \int_0^{v_a} c_a(x)dx + \sum_{i \in N} \sum_{r \in R} w_i^r \tag{24}$$

subject to

$$v_a = \sum_{r \in R} v_a^r, \quad a \in A \tag{25}$$

$$v_a^r \leq f_a w_i^r, \quad a \in A_i^+, i \in N, r \in R \tag{26}$$

$$\sum_{a \in A_i^+} v_a^r - \sum_{a \in A_i^-} v_a^r = g_i^r, \quad i \in N, r \in R \tag{27}$$

$$v_a^r \geq 0, \quad a \in A, r \in R. \tag{28}$$

Spiess and Florian proposed a solution algorithm based on the linear approximation algorithm (Frank and Wolfe, 1956). The application of this algorithm is appropriate since the linear approximation of the objective function (24) results in a problem that is separable by destination r. For each destination, the auxiliary subproblem is equivalent to the problem with constant link times and may therefore be solved efficiently using the label-setting algorithm proposed for the linear cost model. The resulting algorithm may be outlined as follows:

0. (Initialisation) Find a feasible solution (v_0, w_0), where v_0 denotes a feasible vector of total flow v_a and the scalar w^0 denotes the corresponding total waiting time $\sum_{i \in N} \sum_{r \in R} w_i^r$.
   Set $k := 0$.
1. (Subproblem) Set $k := k + 1$.
   Compute $(\tilde{v}, \tilde{w})$ by applying the algorithm of section 2.2 with link times $c_a = c_a\left(v_a^{k-1}\right)$ for each destination $r \in R$.
2. (Line Search) Find λ^k that minimizes the objective function (22) on the line segment
   $(1-\lambda)\left(v^{k-1}, w^{k-1}\right) + \lambda(\tilde{v}, \tilde{w}), 0 \le \lambda \le 1$.
3. (Update) Set $\left(v^k, w^k\right) := (1-\lambda^k)\left(v^{k-1}, w^{k-1}\right) + \lambda^k(\tilde{v}, \tilde{w})$.
   If a stopping criterion is satisfied then STOP.
   Otherwise go to step 1.

3.2 Model of Nguyen and Pallottino - Hyperpaths

Using the hyperpath approach of their uncongested model (section 2.3), Nguyen and Pallottino (1988) assumed that the arc costs $c_{ij}(x)$ are increasing functions of vector flow, modeling the on-board crowding of buses which may affect the passengers' choices. Let y_k denote the flow on hyperpath k and d_l the demand associated with the l-th origin-destination pair. We denote by P_l the set of all hyperpaths connecting the l-th origin-destination pair and by Ω the set of all demand-feasible hyperpath flow vectors

$$\Omega = \left\{ y : \sum_{k \in P_l} y_k = d_l, \quad \text{for all } O/D \text{ pairs } l \right\}.$$

The hyperpath flow vector y induces the arc flow vector $x = \{x_{ij}\}_{(i,j) \in A}$, where

$$x_{ijk} = \alpha_{ijk} y_k$$
$$x_{ij} = \sum_l \sum_{k \in P_l^{ij}} x_{ijk}$$

The authors assumed also that all transit passengers travel on *shortest hyperpaths*, then y^* is an equilibrium flow which satisfies the Wardrop conditions:

$$C_k(y^*) = \min_{k' \in P_l^{ij}} \{C_{k'}(y^*)\} \quad \text{if } y_k^* > 0.$$

As shown by Smith (1979), y^* is a solution to the following variational inequality :

$$(x - x^*)C(x^*) + (y - y^*)w \geq 0, \quad \forall y \in \Omega,$$

in which w is the waiting cost vector with flow independent components

$$w_k = \sum_{i \in N_k} \beta_{ik} w_{ik}, \quad \forall k \in P_l \text{ and } \forall k.$$

Nguyen and Pallottino presented the following basic feasible direction algorithm for this variational inequality :

(1) (Initialisation)
Determine a feasible aggregated arc flow x^1.
Compute the associated total waiting time cost $u^1 = y^1 w$,
set $i := 1$.

(2) (Testing the current flow)
Determine a minimum cost aggregated flow z and the associated waiting cost v with respect to the current cost $c(x^i)$.
If $(z - x^i)c(x^i) + (v - u^i) \geq 0$
then stop (x^i is an equilibrium flow)
else go to step 3.

(3) (Computing an improving direction)
Determine an improving direction $(\tilde{z}, \tilde{v})$ using Frank and Wolfe's descent direction.

(4) (Revising the current flow)
Compute an optimal step length θ^* satisfying :
$$\left[\left(\tilde{z} - x^i\right)C\left(x^i + \theta^*\left(\tilde{z} - x^i\right)\right) + \tilde{v} - u^i\right]\left(\theta - \theta^*\right) \quad \forall 0 \leq \theta \leq 1$$
Revise the current flow and total waiting cost:
$x^{i+1} := x^i + \theta^*(\tilde{z} - x^i)$;
$u^{i+1} := u^i + \theta^*(\tilde{v} - u^i)$;
$i := i + 1$.
Return to step 2.

where the step 2 reduces to the computation of shortest hyperpaths, which can be computed by the label-correcting procedure described in section (2.3). This algorithm is very similar to that proposed by Spiess and Florian (1989).

3.3 Model of De Cea and Fernández – Transit routes

De Cea and Fernández (1993) proposed an extension of their mathematical model (section 2.4) to congested transit networks. As the uncongested case, passengers use *transit routes* to travel from their origin to their destination. The authors assumed that congestion affects waiting times but not in-vehicle travel times. Waiting times at transit stops depend on the total capacity of the attractive set of lines considered and on the total number of passengers using that set of lines. Then, the waiting time experienced by a passenger boarding the route section s (portion of a transit route between two consecutive transfer nodes) at its origin node $i(s)$ will depend on:

- V_s the total number of passengers boarding the same route section, at node $i(s)$,
- V_{is}^+, the total number of passengers boarding, at node $i(s)$, all other route sections that use lines contained in route section s, and,
- $\bar{V}_{is}$, the number of passengers boarding all the lines belonging to route section s at a node before $i(s)$ and alighting after $i(s)$.

The flow V_{is}^+ competes with V^s for the residual capacity of the lines belonging to route section s at node $i(s)$, and the flow $\bar{V}_{is}$ reduces that capacity. If we denote by $\tilde{V}_s = V_{is}^+ + \bar{V}_{is}$ the flow that compete with V^s for the same common capacity, then waiting times will increase as $\tilde{V}_s$ and/or V_s approach the route section capacity K_s. Therefore, the cost function of the route section s takes the following general form:

$$c_s = t_s + \frac{\alpha}{f_s} + \beta.\phi_s(\frac{V_s + \tilde{V}_s}{K_s}),$$

where t_s is the constant in-vehicle travel time, f_s is the total frequency on route section s and α, β are calibration parameters. The third term on this form takes explicitly into account the effect of congestion on waiting times and the form of function $\phi_s(V)$ should be such that c_s is strictly monotone in V_s.

When there is no congestion, waiting time for passengers boarding route section s can be expressed as:

$$\overline{w}_s = \frac{1}{\sum_{l \in \overline{A}_s} f_l},$$

with $\overline{A}_s$ the set of attractive lines associated to route section s and f_l the *nominal frequency* of line l. However, when congestion exists a fraction of the buses arriving to node $i(s)$ will be full. In order to represent this phenomena, De Cea and Fernández introduced the concept of *effective frequency*. If we define an equivalent average waiting time index for each line arriving to node $i(s)$:

$$w_l^i = \frac{\alpha_l}{f_s} + \phi_l(\frac{\tilde{v}_{il}}{k_l}),$$

where ϕ_l must be a monotonically increasing function of $\tilde{v}_{il}$ and k_l is the capacity of a line l, then the effective frequency of line l at stop $i(s)$ is defined as:

$$f_l^{'s} = \alpha_l / w_l^i.$$

The number of passengers using each line in a given route section will be proportional to its relative frequency:

$$v_l^s = x_l^s(v).\ f_l^{'s}.w^s(v).V^s, \tag{29}$$

where integer variables x_l^s depend on the line sections flow vector v and $w^s = 1/\sum_{l \in \overline{A}_s} f_l^{'s}$ is the average waiting time for a passenger using route section s.

While this model uses the effective frequencies to consider congestion over the transit network, it has some important drawbacks. Firstly, this model does not guarantee that lines will not be overloaded since the waiting time functions resemble the delay functions used in the modeling of car networks. Secondly, the authors solve the linear version of the equation (29) and then obtain the same results as those obtained in the uncongested model. Thirdly, they proposed without proof an approach for solving the nonlinear case which does not guarantee the Wardrop equilibrium conditions.

3.4 Model of Wu *et al.* – Hyperpaths

Wu *et al.* (1994) proposed an extension of the nonlinear model of Spiess and Florian (section 3.1) to asymmetric cost functions, which models both the waiting and in vehicle costs as function of the transit flows. It is based on the hyperpath concept on an expanded network, where each arc is assigned a frequency and a travel cost function. The modeling of the waiting cost differs from previous approaches to the problem. It is considered to be the sum of two

terms: one term is the waiting time under uncongested conditions, which depends only on the line headways, and the other term is a cost associated to wait arcs, which is flow dependent.

Let $G=(N,A)$ the general network that is constructed for the model formulation. The nodes of this general network correspond to centroids and transit stops including transfer nodes; the arcs correspond to the segments of the transit lines (in-vehicle arcs), walk arcs, wait arcs and transfer/alight arcs. To each arc $a\in A\setminus A^w$ is associated a cost $s_a(v)$, where A^w denotes the set of wait arcs. The wait arcs are assigned the frequency f_a of the line to which they are connected and all other arcs have infinite frequencies. Wu *et al.* assumed that all transit passengers travel on *shortest hyperpaths*. The cost of hyperpath k associated with a hyperpath flow vector y is defined as:

$$C_k(y)=\sum_{a\in A}\alpha_a^k\left[s_a(v)+W_{ak}+w_a(v)\right],$$

where α_{ijk} and W_{ak} are respectively the arc probabilities and constant arc waiting times associated with the hyperpath k (see section 3.2) and $w_{ak}(v)$ are waiting times added on the wait arcs $a\in A^w$ and are caused by congestion, i.e.,

$$w_a(v)=\begin{cases}\geq 0, & \text{if } a\in A^W\\ =0, & \text{otherwise}\end{cases}$$

As the model of Spiess and Florian, Wu *et al.* assume that the passenger flows on the general network satisfy Wardrop's user-optimal equilibrium conditions and can be alternatively stated as the following variational inequality problem: find $y^*\in\Omega$ such that:

$$s(v^*)^T(v-v^*)+w(v^*)^T(v-v^*)+W^T(y-y^*),\quad \forall v\in\Omega_v \text{ and } \forall y\in\Omega$$

where y^* and v^* are the equilibrium hyperpath and arc flows respectively; Ω and Ω_v are the feasible hyperpath flow and arc flow sets; and W is a vector of waiting times for all the hyperpaths k. Under the continuity and the strict monotony of the arc function $v(\cdot)+w(\cdot)$, the authors provided existence and uniqueness proofs of an equilibrium solution.

Wu *et al.* proposed a symmetric linearization method to solve the transit equilibrium problem. At each iteration l, the nonlinear mapping C is approximated by a symmetric linear mapping:

$$C^l(y)=C(y^l)+\frac{1}{\alpha}B(y^l)(y-y^l) \tag{30}$$

where $B(y^l)$ is a square symmetric and positive matrix and α is a positive parameter. The authors considered only the cases where $B(y^l)$ is the diagonal of the Jacobien matrix of C evaluated at y^l (linearized Jacobi method) or $B(y^l)$ is some fixed matrix B (projection method). In both cases (30) reduces to the solution of the following convex quadratic optimization problem

$$\min_{x\in\Omega}(y-y^l)^T C(y^l)+\frac{1}{2\alpha}(y-y^l)^T B(y^l)(y-y^l). \tag{31}$$

If we define the gap function $GAP(y)$ as $GAP(y)=\min_{x\in\Omega} C(y)^T(y-x)$, we obtain the following symmetric linearization method:

Step 1 (initialization) :

Let ε be a tolerance parameter.

Let K_i^0 be a subset of K_i (set of hyperpaths for the origin-destination pair i).

Let y^0 and v^0 be an initial feasible solution.

Set $l=1$.

Step 2 (shortest hyperpath computation) :

For every $i\in I$ find, based on the current cost $s(v^{l-1})$, compute a shortest hyperpath k_i, and $GAP(y^{l-1})$.

Let $K_i^l = K_i^{l-1}\cup k_i$.

Step 3 (convergence test) :

If $GAP(y^{l-1})\le\varepsilon$ then stop.

Step 4 (S and B computation) :

Based on $s(v^{l-1})$, find $C_k(y^{l-1})$ and $B_k(y^{l-1})$ (if necessary) for $k\in K_i^l$.

Step 5 (projection) :

Solve the quadratic program (31) and let y^l be its solution.

Step 6 (update) :

Compute $v^l=\alpha y^l$ and update $s(v^l)$.

$l=l+1$ and return to Step 2.

The authors proved the global convergence of this algorithm for strongly monotone arc cost mappings. It is be noted that the model of Wu *et al.* (1994) are conceptually equivalent to the one proposed by Nguyen and Pallottino (section 3.2) since the waiting times caused by congestion added on wait arcs are independent from passenger hyperpaths and hence can be viewed as on-board travel times.

3.5 Model of Bouzaïene-Ayari *et al.* – Hyperpaths

Bouzaïene-Ayari (1996) and Bouzaïene-Ayari, Gendreau and Nguyen (1995) proposed a congested assignment model where all travel costs/times and distribution probabilities at bus stops are flow-dependent. To formulate the model, we use the notations in the model of Nguyen and Pallottino (sections 2.3 and 3.2). For any hyperpath assignment $y \in \Omega$, Bouzaïene-Ayari *et al.* assumed that the passenger distribution on the arcs associated with y is v^y where v^y is any fixed point of the arc flow distribution function $\Psi(y,.)$ associated with y, i.e.,

$$v^y = \Psi(y, v^y) \tag{32}$$

where $\Psi(\cdot,\cdot): \Omega \times \Omega_v \to \Omega_v$ is such that, for all $y, v \in \Omega \times \Omega_v$, we have

$$\begin{aligned}
\Psi_{ij}(y,v) &= \sum_l \sum_{k \in P_l} y_k^l \alpha_{ijk}^l(v), \forall (i,j) \in A, \forall v \in \Omega_v, \\
\alpha_{ijk}^l(v) &= \beta_{ik}^l(v) \pi_{ijk}^l(v), \forall (i,j) \in A, \forall v \in \Omega_v, \\
\beta_{jk}^l(v) &= \sum_{i \in N_j^-} \beta_{ik}^l(v) \pi_{ijk}^l(v), \forall j \in N_k - \{r\}, \forall v \in \Omega_v,
\end{aligned}$$

and, for all hyperpath $k \in P_l, \forall l$, the vector of arc distribution probabilities $\pi_k^l(v) = (\pi_{ijk}^l(v))$ associated with v is determined as follows

$$\pi_{ijk}^l(v) = \begin{cases} 1, & \forall (i,j) \in A_k \setminus A_k^m, \\ \dfrac{\left[W_{ij}(v)\right]^{-1}}{\sum_{m \in F_k^+} \left[W_{im}(v)\right]^{-1}}, & \forall (i,j) \in A_k^m, \\ 0, & \text{otherwise.} \end{cases} \tag{33}$$

Equation (33) suggests that the passenger distribution at transit stops is inversely proportional to the line waiting times. For any $v \in \Omega_v$, the hyperpath cost function $F(v) = (F_k^l(v))$ associated with v is computed as follows

$$F_k^l(v) = \sum_{(i,j) \in A} \alpha_{ijk}^l(v) \left[C_{ij}(v) + W_{ijk}^l(v) \right], \ \forall k \in P_l, \forall l,$$

where $C(\cdot) = (C_{ij}(\cdot))$ and $W(\cdot) = (W_{ijk}^l(\cdot))$ are respectively the travel time and waiting time functions over Ω_v. Suppose that the hyperpath flow vector $y^* \in \Omega$ satisfy the Wardrop conditions and the arc vector $y^* \in \Omega_v$ satisfy the fixed point condition (32). Then, y^* and v^* must solve the following equilibrium-fixed point problem:

$$\text{find } (y^*, v^*) \in \Omega \times \Omega_v \text{ s.t.} \begin{cases} v^* & = & \Psi(y^*, v^*) = \alpha(v^*) y^*, \\ F_k^l(v^*) & = & u_l^*, \text{ if } y_k^{l*} > 0, \forall k \in P_l, \forall l \\ u_l^* & = & \min\limits_{k \in P_l} \left\{ F_k^l(v^*) \right\}, \forall l \end{cases}$$

which is equivalent to the following variational inequality-fixed point problem:

$$VIP(v): \text{find } (y^*, v^*) \text{ s.t. } (F(v^*))^T (y^* - y) \le 0, \forall (y, v) \in \Omega \times \Omega_v \left| v = \alpha(v) y. \right. \quad (34)$$

To establish the existence of an equilibrium the authors assume waiting time functions to have uniformly bounded gradients which tend to infinity as the transit flows reach the capacity of the vehicles. Also, the travel time functions are assumed to be strongly monotone, which *prevents* the model being used in the simplest case of constant travel times. Moreover, uniqueness of equilibrium solution is established only for the case in which this strongly monotone term dominates over the waiting time. To compute an equilibrium-fixed point solution, Bouzaïene-Ayari *et al.* proposed the following general solution algorithm based on the successive approximation fixed algorithm:

Step 1: Let ε be a small positive parameter and λ be a sufficiently small relaxation parameter in [0,1]. Set $h \leftarrow 0, y^* \leftarrow y \in \Omega_y$ and $\overline{P}_l^h = \left\{ k \in P_l \left| y_k^{l*} > 0 \right. \right\}, \forall l$.

Set $h \leftarrow h+1$ and $v^h \leftarrow v \in \Omega_v$.

Step 2: Solve problem $VIP(v^h)$ using the linearized Jacobi method.

Step 3: Let (y^{**}, x^{**}) be a solution to $VIP(v^h)$.

If $\left\| v^h - x^{**} \right\| \le \varepsilon$, then stop and $(y^*, v^*) = (y^{**}, x^{**})$.

Step 4: $v^{h+1} \leftarrow (1-\lambda) v^h + \lambda x^{**}$. $h \leftarrow h+1$. Go to Step 2.

Computational results indicate that the algorithm converges slowly.

4 Transit assignment model-hard capacity constraints

In this section, we review congested transit models where capacities of vehicles are taken into account. In this case, not all passengers waiting at a stop can board the first arriving vehicle and part of these passengers have to wait for one on the following vehicles would have available seats in order to board.

4.1 Model of Cominetti and Correa – Common lines

Cominetti and Correa (2001) developed a common-line model with congestion, in which transit vehicles have limited capacities. Passengers are assumed to travel according to shortest hyperpaths/strategies. Consider the simplest network consisting of an origin O connected to a destination D by n bus lines $l_i, i \in A = \{i = 1,...,n\}$, each one characterized by an in-vehicle travel time t_i. Chriqui and Robillard (1975) (see section 2.1) neglected congestion and assumed that passengers can effectively board the first incoming bus. However, limited bus capacities may prevent a passenger from boarding the bus that comes first, increasing his waiting time. To model this situation, Cominetti and Correa assumed that each line l_i is characterized by a differentiable effective frequency function $f_i : [0, \bar{v}_i) \to (0, \infty)$, with $f_i'(\cdot) < 0$ and $f_i(v_i) \to 0$ as $v_i \to \bar{v}_i$. The constant $\bar{v}_i$ is called the saturation flow of the line. The chosen strategy s should minimize the expected transit time $T_s(v)$, including the waiting time $1/\sum_{i \in s} f_i(v_i)$ and the expected in-vehicle travel time $\sum_{i \in s} t_i \pi_i^s$, where $\pi_i^s = f_i(v_i)/\sum_{j \in s} f_j(v_j)$ is the probability of boarding line l_i, i.e.

$$T_s(v) = \frac{1 + \sum_{i \in s} t_i f_i(v_i)}{\sum_{i \in s} f_i(v_i)},$$

which is a function of line-flows so the optimal decision of each passenger is affected by the choice of the others. Let $x > 0$ be the demand for the pair $(0, D)$ and denote by S the set of nonempty subsets $s \subset A$ (possible strategies). The flow x splits into flows $y_s \geq 0$ along the strategies so that $x = \sum_{s \in S} y_s$. If we denote $S_i = \{s \in S : i \in s\}$ the set of strategies containing line l_i, then the vector of strategy-flows $y = (y_s) s \in S$ determines a unique vector of line-flows $v = v(y)$ through the system of equations

$$v_i = \sum_{s \in S_i} y_s \frac{f_i(v_i)}{\sum_{j \in s} f_j(v_j)} \qquad \forall i = 1, \cdots n.$$

Using Wardrop's principle, the strategy-flow vector y^* is an equilibrium if and only if all strategies carrying flow are of optimal time:

$$y_s^* > 0 \Rightarrow T_s(v(y^*)) = \tau(v(y^*)),$$

where $\tau(v) = \min_{s \in S} T_s(v)$. Denote by Y_x the set of all equilibrium y's. Cominetti and Correa have shown that Y_x may be fully characterized by an equivalent optimization problem expressed directly in terms of line-flows v. In particular, denoting V_x the set of line-flows $v(y)$ corresponding to all possible equilibria $y \in Y_x$, the authors proved that, under the condition $x \in (0, \sum_{i=1}^{n} \bar{v}_i)$, there exists a constant $\alpha_x \geq 0$ such that $v \in V_x$ if and only if $v \geq 0$ with $\sum_{i=1}^{n} v_i = x$ and

$$\frac{v_i}{f_i(v_i)} \begin{cases} = \alpha_x & \text{if } t_i < \tau(\hat{v}(\alpha_x)) \\ \leq \alpha_x & \text{if } t_i = \tau(\hat{v}(\alpha_x)) \\ = 0 & \text{if } t_i > \tau(\hat{v}(\alpha_x)) \end{cases} \tag{35}$$

where $\hat{v}_i(\alpha)$ denotes the unique solution z of the equation $z / f_i(z) = \alpha$.

Using the common-line model described above, Cominetti and Correa proposed an equilibrium model for a general transit network $G = (N, A)$. Let $i(a)$ and $j(a)$ be respectively the tail and head nodes of $a \in A$ and denoting by $A_i^+ = \{a : i(a) = i\}$ and $A_i^- = \{a : j(a) = i\}$ the sets of arcs leaving and entering node $i \in N$. The set of destinations is denoted $D \subseteq N$, and for each $d \in D$ and every node $i \neq d$ a fixed demand $g_i^d \geq 0$ is given. Also, each arc $a \in A$ is characterized by a continuous travel time function $t_a : V \to [0, \bar{t}_a)$ where $\bar{t}_a$ is a finite upper bound, and an effective frequency function $f_a : V \to [0, \infty)$ where $V = [0, \infty)^{A \times D}$ denotes the space of destination-arc-flow vectors v with $v_a^d \geq 0$ for all $a \in A, d \in D$. A vector $v \in V$ is said to be feasible if $v_a^d = 0$ for all $a \in A_d^+$ (no flow with destination d leaving node d) with each v^d satisfying flow conservation

$$g_i^d + \sum_{a \in A_i^-} v_a^d = \sum_{a \in A_i^+} v_a^d \quad \forall i \neq d.$$

Consider a passenger heading towards destination d and reaching an intermediate node i in his trip. To exit from i he can use the arcs $a \in A_i^+$ to reach the next node $j(a)$. By taking the arc travel times $t_a(v)$ and the transit times $\tau_{j(a)}^d$ from j_a to d is given, the decision faced at node i is a common-line problem with travel times $t_a(v) + \tau_{j(a)}^d$ and effective frequencies corresponding to the services operating on the arcs $a \in A_i^+$. The solution of this common-line

problem determines the transit time τ_i^d from i to d, which can be used recursively to solve the up-stream nodes. Since all the variables τ_i^d and v_a^d must be determined at the same time, the transit network model is stated as a set of simultaneous common-line problems (one for each pair (i,d)) coupled by flow conservation constraints. More precisely, for each $v \in V$ the flow entering node i with destination d is defined as $x_i^d(v) = g_i^d + \sum_{a \in A_i^-} v_a^d$, while the time-to destinations functions $\tau_i^d(v)$ are the unique solution of the generalized Bellman equations:

$$\begin{cases} \tau_d^d = 0 \\ \tau_i^d = \min\limits_{s \subset A_i^+} \dfrac{1 + \sum\limits_{a \in s} \left[t_a(v) + \tau_{j(a)}^d \right] f_a(v)}{\sum\limits_{a \in s} f_a(v)}, \quad \forall i \neq d. \end{cases} \tag{36}$$

Let $V_i^d(v)$ denote the set of equilibrium flows corresponding to a common-line problem defined by the arcs A_i^+, with total flow $x_i^d(v)$, constant travel times $t_a^d(v)$, and diagonal effective frequencies $f_a^d(\cdot)$ obtained from $f_a(\cdot)$ by considering it as a function of the flow v_a^d alone, and keeping the other variables fixed. A feasible flow v is called a network equilibrium if for all $d \in D$ and $i \neq d$ the flows $\left(v_a^d\right)_{a \in A_i^+}$ belong to $V_i^d(v)$. Cominetti et Correa proved that, if for each $d \in D$ and every $i \neq d$ there is a path from i to d in the infinite frequency subgraph of G, then the set of network equilibrium flows V^* is nonempty.

4.2 Model of Cepeda – Node strategies

In his doctorate thesis, Cepeda (2002) extended the common line model of Cominetti and Correa (section 4.1) to general transit networks, which have strict limits on the capacity of transit line segments. The authors used the concept of *node-strategy* to state a new formulation of this problem. For each node-destination (i,d), a node-strategy is a subset of arcs exiting from node i. This definition use the concept 'node-destination pair' instead of 'origin-destination' because each node $i \in N$ must be considered as a potential origin since it may receive transfer flows coming from other nodes and heading towards destination d. A strategy (hyperpath) k can be decomposed into node-strategies, each one for every node $i \in N$. Denoting by Ω_S the set of all node-strategy flows and h_k^d the flow associated with node-strategy $k \subset A_i^+$ heading towards destination d. Let $S_k^d(v)$ be the expected travel time to travel from i to d using node-strategy k and $\tau_i^d(v)$ the expected cost of shortest

hyperpath from i to d (with frequencies $f_a(v)$ and travel times $t_a(v)$) given by the recursive equations (36). The set $\Omega_S \times \Omega_V^D$ of feasible flows contains all vectors $(h,v) \in \Omega_S \times \Omega_V^D$ such that

$$\sum_{a \in A_i^+} v_a^d = g_{id} + \sum_{a \in A_i^-} v_a^d, \quad i \in N, \quad d \in D;$$

$$h_k^d = \sum_{a \in k} \left(h_k^d \frac{f_a(v)}{\sum_{a' \in k} f_{a'}(v)} \right) \quad a \in A, \quad d \in D;$$

$$v_a^d \geq 0, \quad a \in A, \quad d \in D;$$

$$h_k^d \geq 0, \quad k \subset A_i^+, \quad i \in N, \quad d \in D;$$

$$f_a(v) > 0, \quad a \in A;$$

A feasible vector $v \in \Omega_V^D$ is an equilibrium flow if there exists $h \in \Omega_S$ satisfying the Wardrop's conditions

$$S_k^d(v) \begin{cases} = \tau_i^d(v) & \text{if } h_k > 0 \\ \geq \tau_i^d(v) & \text{if } h_k = 0 \end{cases}, \quad k \subset A_i^+, i \in N, d \in D,$$

with $v \in V(h) = \left\{ v \in \Omega_V^D : (h,v) \in \Omega_S \times \Omega_V^D \right\}$; where the cost of the node-strategy $k \in A_i^+$ is given by:

$$S_k^d(v) = \frac{1 + \sum_{a \in k} f_a(v)(t_a(v) + \tau_{j(a)}^d(v))}{\sum_{a \in k} f_a(v)}.$$

As a direct application of the common-line equilibrium conditions (35) of section 4.1, Cepeda proved that $v \in \Omega_V^D$ is a network equilibrium flow if there exist number $\alpha_i^d \geq 0$ such that for all $d \in D$ and $i \neq d$

$$\frac{v_a^d}{f_a(v)} \begin{cases} = \alpha_i^d & \text{if } t_a^d(v) < \tau_i^d(v) \\ \leq \alpha_i^d & \text{if } t_a^d(v) = \tau_i^d(v) \\ = 0 & \text{if } t_a^d(v) > \tau_i^d(v) \end{cases}$$

where $t_a^d(v) = t_a(v) + \tau_{j(a)}^d(v)$. This characterization allows to derive a gap function with the property that each of its minimizers corresponds to a network equilibrium. More precisely, Cepeda proved that for all $v \in \Omega_V^D$, $d \in D$ and $i \neq d$ the inequality

$$\sum_{a \in A_i^+} \left[t_a(v) + \tau_{j(a)}^d(v) \right] v_a^d + \max_{a \in A_i^+} \frac{v_a^d}{f_a(v)} \geq \tau_i^d(v) \sum_{a \in A_i^+} v_a^d$$

holds. Moreover $v \in \Omega_v^D$ is an equilibrium if all these inequalities are satisfies as equalities. This result shows that the equilibrium flows v are the global minima of the function

$$\sum_{d\in D, i\neq d}\left[\sum_{a\in A_i^+}\left[t_a(v)+\tau^d_{j(a)}(v)\right]v_a^d+\max_{a\in A_i^+}\frac{v_a^d}{f_a(v)}-\tau_i^d(v)\sum_{a\in A_i^+}v_a^d\right]$$

with optimal value equal to 0. A trivial computation permits to state the problem of finding equilibrium flows in the form

$$(P)\qquad \min_{v\in\Omega_v^D} G(v)$$

where $G(v)=\sum_{d\in D}\left[\sum_{a\in A}\left[t_a(v)\right]v_a^d+\sum_{i\neq d}\max_{a\in A_i^+}\frac{v_a^d}{f_a(v)}-\sum_{i\neq d}g_i^d\tau_i^d(v)\right]$. Although problem (P) characterizes the network equilibria, if one attempts to solve it by using a descent method, the iterates may be trapped into a local minimum which is not an equilibrium. Cepeda proved that all local minima of (P) are global, but only for the special case of common-lines. However it is unlikely that such a result holds in general. Now, since the optimal value of (P) is known, an approach to solve it is to use a heuristic method and to evaluate the deviation from optimality of the computed solutions by using the value of the objective function. At each iteration, this method computes a transit network equilibrium for the linear network obtained by fixing the travel times and the frequencies at the values determined by the current flows, and then updates these flows by averaging the previous iterate and the newly computed solution. An equilibrium for the linear cost network can be found (Spiess, 1983) by solving

$$\min_{v\in\Omega_v^D}\sum_{d\in D}\left[\sum_{a\in A}\left[t_a(v)\right]v_a^d+\sum_{i\neq d}w_i^d\right]$$

$$\text{s.t. } v_a^d\le w_i^d f_a \quad d\in D, i\neq d, a\in A_i^+$$

which amounts to compute a shortest hyperpath for each destination $d\in D$, and then determining the corresponding destination-arc-flow vector. Choosing $\alpha_k\in(0,1)$ with $\alpha_k\to 0$ and $\sum_{k=0}^{\infty}\alpha_k=\infty$, this method can be described as follows:

Initialize: find $v^0\in\Omega_v^D$ and set $k\to 0$

while: $G(v^k)>\varepsilon$

- compute $t_a=t_a(v^k)$ and $f_a=f_a(v^k)$
- compute shortest hyperpaths for each $d\in D$

- determine the induced flows $\hat{v}_a^d$
- update $v^{k+1} = (1-\alpha_k)v^k + \alpha_k \hat{v}$
- set $k \leftarrow k+1$

end

Stop: v^k is a solution with gap $G(v^k) \leq \varepsilon$.

Although the numerical results obtained with this algorithm are satisfactory and provide acceptable values of the gap function in a reasonable number of iterations, the empirical convergence of the algorithm raises the interesting issue of devising a rigorous convergence proof.

4.3 Model of Hamdouch *et al.* – Travel strategies

Hamdouch *et al.* (2002) developed a strategic equilibrium model with fixed schedules and rigid capacities where the probability of accessing a given bus line is not related to the arrival process but with the queuing process that takes place at boarding nodes. In this context, a strategy consists of an *ordered* subset of available lines, and users board the first vehicle from that subset that has an available seat. For a given user, the boarding probability clearly depends both on the residual capacity and the position of the user within the queue. Such a model was proposed by Marcotte, Nguyen and Schoeb (2000) and then extended by Hamdouch (2002) and Hamdouch *et al.* (2002) to priority networks. As an example, let us consider the graph illustrated in Figure 7, where we associate with each arc a delay (cost) and a capacity (parenthesized). The subpaths 1-3 and 2-3-5 correspond to two bus lines *A* and *B* with rigid capacities equal to 10, and the remaining arcs can be interpreted as walking arcs, i.e., arcs with infinite capacities. Demand for the sole origin-destination pair (1,5) is set to 15.

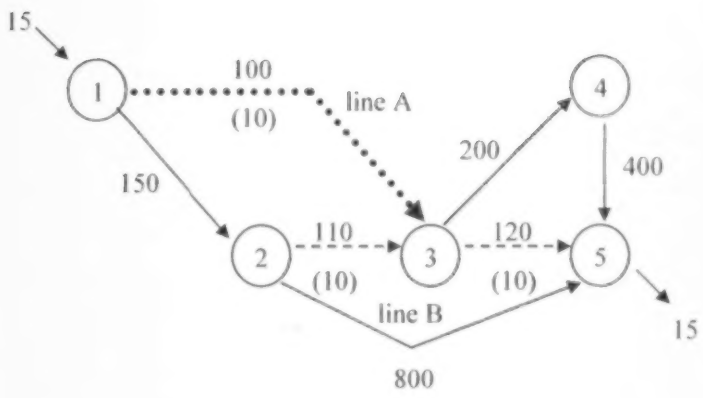

Figure 7: An example of a network with rigid capacities

User behaviour is dictated by a strategy that specifies, at each node of the network, an ordered set of adjacent nodes. In our example, the only two strategies of interest are shown in Table 2.

nodes :	1	2	3	4	5
s_1 :	[2]	[3,5]	[5,4]	[5]	[ ]
s_2 :	[3,2]	[3,5]	[5,4]	[5]	[ ]

Table 2: The two strategies of interest for the network of Figure 7

It is clear from the network topology that users boarding a bus of line B at node 2 and travelling on arc (2,3) can always access arc (3,5) by staying on board, and have therefore a natural priority over competing users. In the model of Hamdouch *et al.*, a strategic equilibrium is found according to this natural priority in loading passengers to segments of transit lines. The authors made the assumption that if u_{jk} denotes the capacity of arc (j,k), the probability π_{jk} if accessing arc (j,k) for each of x competing users is given by $\pi = \min\{1, u_a / x\}$. These access probabilities π^s_{jk} allow to compute the strategic expected costs C^s:

$$C^s = \sum_{p \in P^s} \prod_{(j,k) \in p)} \pi^s_{jk} \Big(\sum_{(j,k) \in p)} c_{jk} \Big),$$

where P^s denotes the set of all elementary paths including in the strategy s and c_{jk} is the travel cost of arc (j,k). A strategic equilibrium is reached when all users are assigned to

strategies of minimal expected costs. It is not difficult to check on this small example that the unique equilibrium satisfying the priority loading is $x^* = (x_1^*, x_2^*) = (15,0)$. The expected cost of strategy s_1 is equal to 570, which is less than 700, the expected cost of the inactive strategy s_2, hence the equilibrium conditions are satisfied.

In general, a vector x of strategic flows is a strategic equilibrium if and only if it is demand-feasible and satisfies the variational inequality

$$\langle C(x), x - y \rangle \leq 0 \tag{37}$$

for every demand-feasible strategic vector y. The evaluation of the cost function C requires the knowledge of access probabilities on the arcs of the network. Hamdouch *et al.* developed a nontrivial priority loading procedure, which computes these access probabilities and the expected costs of strategies. The authors proved that the cost mapping induced by the loading process is lower semi-continuous as function of incoming flows. This property ensures the existence of at least one equilibrium solution.

Hamdouch *et al.* proposed a solution algorithm for solving the variational inequality (37). This algorithm, which mimics the method of successive averages (MSA), generates a sequence of iterates by mixing the current strategic flow with the flows corresponding to optimal strategies. Next, the loading procedure yields updated link flows, access probabilities and expected cost vector $C(x^{k+1})$, allowing the continuation of the process. In mathematical terms, this is expressed as

$$x^{k+1} = (1 - \theta^k) x^k + \theta^k \bar{x}(x^k), k = 1, 2, \ldots$$

where $\theta^k \in (0,1)$ and

$$\bar{x}_s^l(x^k) = \begin{cases} d_{qr} & \text{if } s = s^* \\ 0 & \text{otherwise} \end{cases}, \ \forall (q,r),$$

where d_{qr} denotes the demand for pair (q,r). The construction of an optimal strategy s^* is based on the recursive process developed by Nguyen and Pallottino (1988). At node j, let $E_j^{s^*}$ be the optimal preference order at node j. The expected travel costs $\omega_j^{s^*}$ from node j to the destination r are then computed by scanning the nodes in reverse topological order, and applying the recursion

$$w_j^{s^*} = \begin{cases} \infty & \text{if } j > r \\ 0 & \text{if } j = r \\ \sum_{k \in E_j^{s^*}} \pi_{jk}^{s^*}(c_{jk} + \omega_k^{s^*}) & \text{if } j < r \end{cases}$$

For improved convergence of the algorithm, the step size was chosen adaptively and independently for each OD pair. More precisely:

$$\theta_s = 1 - \frac{C^{s^*}}{C^s}.$$

The algorithm was halted as soon as the relative gap function, a measure of disequilibrium, defined as

$$g(x) = \frac{\max_{y \in X} \langle C(x), x - y \rangle}{\langle C(x), x \rangle}$$

became smaller than some predetermined tolerance.

5 Conclusion

This review article has presented several approaches for handing capacity constraints in a transit route choice model. In practice, the methods proposed by Spiess and Florian (1989), and Cepeda, Cominetti and Florian (2003) have been used successfully. The model of De Cea and Fernández (1993) has also been incorporated in a method that is used in practice. While the issue of capacity in transit assignment is an important one, this issue may still attract new contributions.

References

Andreasson, I. (1976) A method for the analysis of transit networks, *Proceedings of the 2nd European Congress on Operations Research* (Edited by Mark Roubens), North Holland, Amsterdam.

Bouzaïene-Ayari, B., Gendreau, M. and Nguyen, S. (1995) An equilibrium-fixed point model for passenger assignment in congested transit networks, Publication CRT-95-57, Centre for Research on Transportation, Université de Montréal.

Bouzaïene-Ayari, B. (1996) Affectation statique des passagers dans les réseaux des transport en commun: modélisation et algorithme de résolution, Ph.D. dissertation, Département d'informatique et de recherche opérationnelle, Université de Montréal.

Cepeda, M. (2002) Modèle d'équilibre dans les réseaux de transport en commun : le cas des capacités explicites des services, Ph.D. dissertation, Département d'informatique et de recherche opérationnelle, Université de Montréal.

Cepeda, M., Cominetti, R. and Florian, M. (2003) A transit network model with strict capacity constraints: characterization and computation of equilibria, Publication CRT-2003-15, Centre for Research on Transportation, Université de Montréal.

Chapleau, R. (1974) Réseaux de transport en commun : structure informatique et affectation, Ph.D. dissertation, Département d'informatique et de recherche opérationnelle, Université de Montréal.

Chriqui, C. and Robillard, P. (1975) Common bus line, *Transportation Science*, **9**, 115-121.

Cominetti, R. and Correa, J. (2001), Common-lines and passenger assignment in congested transit networks, *Transportation Science*, **35**, 250-267.

De Cea, J., Bunster, J.P., Zubieta, L. and Florian, M. (1988) Optimal strategies and optimal routes in public transit assignment models, *Traffic Engineering and Control*, **29** (10), 520-526.

De Cea, J. and Fernández, E.(1989) Transit assignment to minimal routes: An efficient new algorithm, *Traffic Engineering and Control*, **30**, 491-494.

De Cea, J. and Fernández, E. (1993) Transit assignment for congested public transportation systems: An equilibrium model, *Transportation Science*, **27**, 133-147.

Dial, R.B. (1967) Transit pathfinder algorithm, *Highway Research Record*, **205**, 67-85.

Fearnside, K. and Draper, D.P. (1971) Public transport assignment – A New approach, *Traffic Engineering and Control*, **13**, 298-299.

Frank, M. and Wolfe, P. (1956) An algorithm for quadratic programming, *Naval Research Logistics Quarterly*, **3**, 95-110.

Gendreau, M. (1984) Une étude approfondie d'un modèle d'équilibre pour l'affectation des passagers dans les réseaux de transport en commun, Ph.D. dissertation, Département d'informatique et de recherche opérationnelle, Université de Montréal.

Hamdouch, Y. (2002) Affectation statique et dynamique des usagers dans un réseau de transport avec capacités rigides, Ph.D. dissertation, Département d'informatique et de recherche opérationnelle, Université de Montréal.

Hamdouch, Y., Marcotte, P. and Nguyen, S. (2002) Capacitated traffic assignment with priority rules, Publication CRT-2002-41, Centre for Research on Transportation, Université de Montréal.

Hasselström, D. (1981) Public transportation planning – A mathematical programming Approach, Ph.D. dissertation, University of Gothenburg, Sweden.

Lampkin,W. and Saalmans, P.D. (1967) The design of routes, service frequencies and schedules for municipal bus undertaking: A case study, *Operations Research Quarterly*, **18**, 375-397.

Last, A. and Leak, S.E. (1976) Transept: A bus model, *Traffic Engineering and Control*, **17**, 14-17, 20.

Le Clercq, F. (1972) A public transport assignment method, *Traffic Engineering and Control*, **14**, 91-96.

Mandl, C.E. (1980) Evaluation and optimization of urban public transportation networks, *European Journal of Operational Research*, **5**, 396-404.

Marcotte, P. and Nguyen, S. (1998) Hyperpath formulations of traffic assignment problems, *Equilibrium and Advanced Transportation Modelling* (P. Marcotte and S. Nguyen, eds), Kluwer Academic Publisher, 175-199.

Marcotte, P., Nguyen, S. and Schoeb, A. (2000) A strategic flow model of traffic assignment in capacitated networks, Publication 2000-10, Centre for Research on Transportation, Université de Montréal.

Nguyen, S. and Pallottino, S. (1988) Equilibrium traffic assignment for large scale transit networks, *European Journal of Operational Research*, **37**, 176-186.

Nguyen, S., Pallottino, S. and Malucelli, P. (2001) A modeling framework for the passenger assignment on a transport network with time-tables, *Transportation Science*, **55B**, 238-249.

Rapp, M.H., Mattenberger, P., Piguet, S. and Grandpierre, A.R. (1976) Interactive graphic system for transit route optimization, *Transportation Research Record*, **559**, 73-88.

Smith, M.J. (1979) The existence, uniqueness and stability of traffic equilibria, *Transportation Research*, **13B**, 295-304.

Spiess, H. (1983) On optimal route choice strategies in transit networks, Publication 286, Centre for Research on Transportation, Université de Montréal.

Spiess, H. and Florian,M. (1989) Optimal strategies : a new assignment model for transit networks, *Transportation Research*, **23B**, 83-102.

Wardrop, J.G. (1952) Some theoretical aspects of road traffic research, *Proceedings of the Institution of Civil Engineers, Part II*, 325-378.

Wu, J.H. and Florian, M. (1993) A simplicial decomposition method for the transit equilibrium assignment problem, *Annals of Operations Research*, **44**, 245-260.

Wu, J.H., Florian, M. and Marcotte, P. (1994) Transit equilibrium assignment: a model and solutions algorithms, *Transportation Science*, **28**, 193-203.